Journal of Critical Reviews HYBRID REDUCTION ALGORITHM WITH CAT SWARM OPTIMIZATION FOR CHURN PREDICTION

Article · May 2020

DOI: 10.31838/jcr.07.07.130

Sujatha Srinivasan
SRM Institute for Training and Development, Chennai, India

Suseendran G.
VELS INSTITUTE OF SCIENCE, TECHNOLOGY & ADVANCED STUDIES (VISTAS), CHENNAI.

Journal of Critical Reviews

ISSN-2394-5125 Vol 7, Issue 7, 2020

HYBRID REDUCTION ALGORITHM WITH CAT SWARM OPTIMIZATION FOR CHURN PREDICTION

[1]J.S.T.M. Poovarasi, [2] Sujatha Srinivasan, [3] G.Suseendran

[1] Research Scholar, School Of Computing Science, Vels Institute Of Science, Technology And Advanced Studies, (VISTAS) Chennai, Tamil Nadu, India,
Email:Jpoovarasi@Gmail.Com
[2] Associate Professor, Department Of Computer Science And Applications, SRM Institute For Training And Development, Chennai, Tamil Nadu, India,
Email: Ashoksuja08@Gmail.Com
[3] Assistant Professor, School Of Computing Sciences, Vels Institute Of Science, Technology And Advanced Studies,(VISTAS), Chennai.
Email:Suseendar_1234@Yahoo.Co.In

Received: 11.02.2020 ***Revised: 04.03.2020*** ***Accepted: 01.04.2020***

Abstract
Customer complexity is a main issue and for large companies is the main problem. Considering the immediate impact on firms 'earnings, companies are trying to change strategies to calculate customer concerns. Consequently, it is very important to find a way to solve this problem by differentiating the factors that increase the client's depression. The chief involvement of this study is to progress an effective churn prediction prototypical using a hybrid approach. Here, initially, data is collected from the dataset and the missing data is removed at the pre-processing stage. Then, to reduce the problem, the input dataset is enhanced as a dimension reduction function. For dimensional reduction, the proposed method uses a hybrid technique. Here, PCA and LDA algorithm are hybridized to reduce dimensionality. After the dimensionality reduction process, the reduced dataset is provided to the optimal continuous neural network (ORNN). Here, the traditional RNA classifier is trained with Cat Swarm Optimization (CSO). In this work, Tera Data Center at Duke University churn set of predictive data for the calculation, the measured performance. Finally, the performance of the proposed model is estimated at different scales, and it is recognized that the proposed system, designed with dimensional reduction through optimal classification methods, performs better with 95.08% classification accuracy compared to other classification models.

Keywords:- Customer churn, Principal Component Analysis, Linear Discriminant Analysis, optimal recurrent neural network, Cat Swarm Optimization, prediction, dimension reduction.

DOI: http://dx.doi.org/10.31838/jcr.07.07.130

INTRODUCTION

Current organizations rely on the customer as a significant resource, and thus the ability to improve their ability to retain customers is a significant concern [1]. Therefore, the need for a model to keep existing customers in order to improve the business establishment obtains customer protection, which is a costly endeavor [2]. Various research works have the potential to analyze customer ID and encourage customers to accelerate when focusing on advertising efforts [3] [4]. Research contracts with early identifiable sources of potential problems that control the customer's pulse, resulting in profit for the retail location or business, including the sale of goods [5]. Customer Relationship Management (CRM) is a significant component of an association. A significant issue in any client-related association is keeping customers [6]. This is due to the frustration experienced by the market and the high concentration of customers. Customer care is a significant requirement for the efficient functioning of an association. Getting new customers is estimated to be five to several times more expensive [7].

Loss of customers can result in loss of income misfortune and brand credibility and trigger the company's decision [8]. Customer Dilemmas / Scraped Area is the option of leaving a client with an association [9]. The significant problem with churn prediction is that there are a few motives behind a customer. Differentiating these factors is complicated because the causes are not immediate [10]. They rely on customers who are close to home perspectives and products or the management used by the company's customers [11]. Associations must be predicted before this occurs. The biggest customer can access information and make predictions, and in any case, the idea of information is the biggest disadvantage of the forecasting tool. Focused marketing efforts often rely on customer-generated benefits [12]. This information reflects the actual personal behavior of customers. Customer requirements determine management roles and influence the hierarchical structure of the company to focus on specific types of customers [14]. Customer complexity is one of the toughest problems due to the significant increase in the media transmission segment [15].

Because of the significant cost associated with this, rather than contracting new people, the focal point of professional partnerships is retaining existing customers. The client expects to be able to identify a potential problem before leaving the company, and the company's enthusiasm has expanded significantly over the years. Along these lines, a few techniques, for example, particle swarm optimization, genetic algorithm, data processing, neural network and machine learning have been used. This test eliminates cat swarm optimization in Bjorn prediction, [16] [17].

LITERATURE REVIEW

A unique investigation method for client problem prediction was developed by Verbeke, W.,et al. [18] In 2014. This test analyzes the use of each other's data for the client's depression prognosis. An innovative way to integrate non-Markovian organizational influences with social classifiers and to provide an innovative parallel visual arrangement to join the social and non-social order model. The effects of two actual situational investigations on large-scale telco information indices were visualized; Contains both organized (call log records) and unorganized (customer related) data around numerous approvals.

The advantage of increasing the strategic model for customer problem forecasting using general calculations has been explored by Stripling, E., et al. [19] In 2018. This practice seeks to develop profitable churn models for the maintenance crusade to meet the business prerequisite for maximizing benefits. In a large study of nine true information codes, Prof Logit shows the most common, best-in-class EMPC executions for general excellence, benefit-based accuracy and review. Because of the similarity of the rope, Prof Logit additionally expresses benefit-based component determinations, in which features are selected, which in some way or another are precluded by precisely based action, which is another compelling finding.

The initial sad prediction of personalized attention in small social games was divided by Milosevic, M., et al. [20] in 2017. For an early prediction, we looked at the creation of basic machine learning models and the use of information from Top Leven's 2,000,000 players - an online versatile game as football manager. To avoid the problem, they monitor client activity, recognize the game highlights that attract the client, and then use that information to retain customized message pop-ups with a reason to pull customers into the game. In this strategy, they had the opportunity to reduce it to 28%, which is attributed to the large number of customers, which has a positive and positive effect on the business.

In 2018 Hopper, S., et al. [21] Zorn has developed a decision tree that is beneficial for prediction. The largest benefit quantity for EMPC has been established, thus selecting the most beneficial churn model. They present another classifier that legally incorporates the EMPC metric in model development. This strategy, called Prof Tree, uses growth calculation to identify benefit-based selection trees. A large review of the actual datasets of different media transmission specialist systems shows that the professor has made tremendous improvements to more precisely implemented woodworking techniques.

Lu, N., and others have developed a predictive paradigm in the use of privilege in the telecommunications industry. [22] In 2014. This study attempts to isolate customers into two incremental calculations. Thus, a high-risk client cluster is recognized. Strategic repetition was used as an exemplary student in this choice, and a Churn prediction model was developed for each group separately.

Idris, A., Et al., Developed a predictive predictive framework for telecommunications using channel wrapper and group characterization. [23] In 2016. In this paper, a particle swarm optimization system has been used. Developed to integrate learning and successful filter-wrapper classifiers. In the long run, Random Forest, Rotation Forest, RotBoost, and SVM have been working on this.

Castro, E. Developed by G, et al. The classifiers are then assessed and used as predictive activation metrics. One of the techniques, time-tested flight field testing, proved acceptable outcomes, with the option of providing maintenance crucifixion results 20% higher than the RFM method.

Amin, A., et al. [25] In 2018, the media exchange industry was shattered by a customer's sad prediction. A genetic computation and the NN have been utilized. The dataset can be assembled into various zones depending on the separation factor, which then falls into two classifications; Customers expect (i) high trust and (ii) low confidence. The study results show the best performance of accuracy, review and accuracy.

Proposed model for customer churn prediction:
The main purpose of the proposed method is to predict the customer churn (CC) based on a hybrid dimension reduction approach with optimal classification. The proposed Customer churn Prediction (CCP) is given in Model 1. The proposed approach has three main phases, namely pre-processing, dimensional reduction and prediction. Initially, processed data is pre-processed to obtain good predictive accuracy. Then, dimensional reduction is done with the help of a hybrid approach called PCLDA. Then, the dimensionally reduced dataset prediction condition is presented. This phase has two stages - training and testing. During the training phase, the Dimension Reduced Database is trained with the help of an RN classifier. Weight beliefs are improved through the assistance of cat swarm optimization (CSO) to recover RNA classification. In the test process, the given test samples are classified as churn or non-churn data.

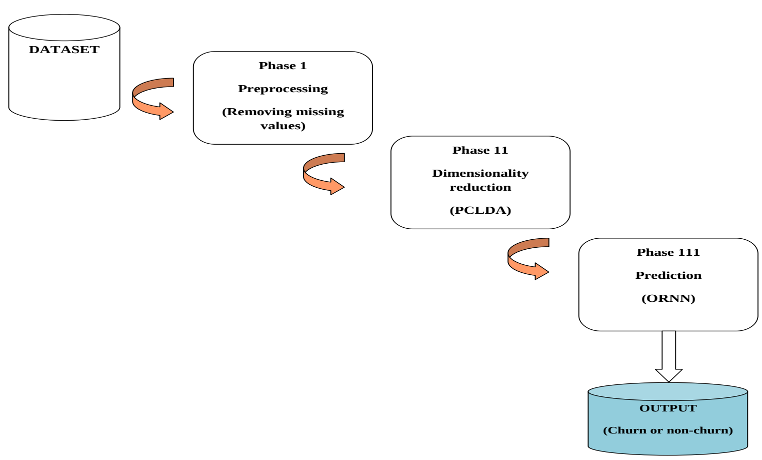


Figure 1: The proposed churn prediction model

Removing missing values in the preprocessing stage
Pre-implementation is an important step;this includes dealing with missing values. In addition to missing values, the data are characterized by their inaccuracy, inconsistency, redundancy, spacing, and invalidity.

Dimensionality reduction using hybrid PCLDA
Dimension reduction is an important process for classification, recognition and prognosis. This is because many features are a major obstacle to the classification process. To overcome the problem, dimensional reduction is introduced. For dimensional reduction, hybridized PCLDA is used in this paper. PCLDA is a combination of PCA and LDA. The operational principle of dimensionality reduction is explained below;

Principal component analysis (PCA):
PCA is one of the standard unsupervised techniques to decrease the dimensionality of the dataset lackingdefeated important

information in the original dataset. Consider training data for input I and system conditions, and use S to estimate the confirmatory output. This can be seen in the PCA transformation by subtracting the MSE. To see this, let there be a feature vector $i \in R^d$, a reduced dimensional aspect vector $y \in R^h$, and a reconstructed feature vector $\hat{i} \in R^d$. Then write the MSE as follows;

$$MSE = E\left[\left\|i - \hat{i}\right\|^2\right] \qquad (1)$$

Where;

$E[\bullet]$ →Expectation operation with respect to i

$\|\bullet\|^2$ → Norm squared value

Basically, PCA transformation Φ is utilized to minimize the size of the dimension since d-dimension space to h-dimensional feature space. The size of Φ is $d \times h$. The PCA transformation can be written as follows;

$$\Phi : i \to y \quad or \quad y = \Phi^T i \qquad (2)$$

The Eigen value problem related to Φ can be considered as surveys;

$$\sum_i \phi_j = \lambda_j \phi_j \qquad (3)$$

Where $\Phi = \{\phi_j : j = 1,2,\ldots,h\}, \ \phi_j \in R^d$ and $\sum_i$ is the covariance matrix of all input d-dimensional vectors. The expression λ_j represents Eigen values conforming to ϕ_j. The eigenvectors $(\phi_1, \phi_2, \ldots, \phi_h)$ of Φ would be decided like that their corresponding Eigen values or are $\lambda_1 > \lambda_2 > \ldots > \lambda_h$.

Linear Discriminant Analysis:

Linear discriminant analysis is known as Fisher linear discriminant analysis and it also classifies classes. This method can be used mainly for multi-class problem. It can also reduce the dimension between the class-scatter matrix and the class-scatter matrix. At LDA the project ranges from the H-dimensional feature space to the h dimensional space, where the patterns of the classes of $k < h$ models are detached[26].

The transformation for a c-Class problematic can write as follows;

$$s = W^T y \qquad (4)$$

Where, $s \in R^k$ and y is selected since equation (2).By increasing the transformation matrix W, Fisher's criterion is calculated.

$$J(W) = \left|W^T S_B W\right| / \left|W^T S_W W\right|$$

(5)

The computation of connecting-class scatters metric S_B with class scatter matrix S_W can be computed from the vectors y. The transformation matrix W is given by [4]:

$$S_B w_i = \lambda_i S_W w_i \qquad (6)$$

Where, $W = \{w_i : i = 1,2,\ldots,k\}$. The eigenvectors w_i (columns of W) correspond to the Eigen values λ_i.

Hybrid PCLDA for feature reduction:

The final dimension reduction is done in this section. Initially, training data is provided to the PCA to obtain reduced data space. Finally, a feature vector y is obtained. Then, the reduced feature vector y is given to the LDA, which produces discriminatory features S. The gradual process of feature vector reduction is given below;

Step 1:consider the input vector I and evaluate the PCA transmission using equation 3.

Step 2: Development of the data in the subordinate dimensional space using Equation 2. This gives the 'y' feature vectors.

Step 3: Evaluate transformation W using Eigen value decomposition (equation 6).

Step 4: Plot the 'y' feature vectors on to k-dimensional space by using equation 5. This will give s feature vectors.

Prediction using ORNN

After the dimension reduction process, the dimensional reduced dataset is provided to input the ORNN classifier to classify the data as semi-spiral or non-spiral data. RNN is a special type of NN that is primarily used for classification, prognosis and recognition. RNN contain two phases and consists of three layers, for example, the level of training and testing stage and the input layer, hidden layer and output layer. The hidden layer contains the hidden and contextual layers. One step delays in the feedback path, so that the topology is identical to the feed-forward network, but the concept of the hidden layer, the results is used feedback signals. Weight principles are improved through the benefit of the CSO algorithm to improve the RNN. The basic structure of the RNN is given in Figure 2.

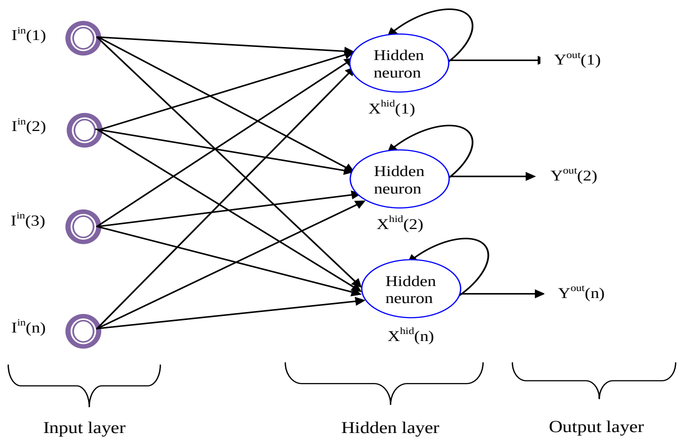


Figure 1: Basic diagram of the recurrent neural network

RNN is trained using training samples $(a_1, b_1), (a_2, b_2) \ldots (a_n, b_n)$ with $a_i \in R^n$, $b_i \in R^n$ for $1 \le i \le n$. Here, a_i denotes the input vector and b_i denotes the target output.

Step 1: Initially, the dimensional reduced dataset is provided for the input of the RNN classifier and assigns their optimal weights.

Step 2: Commonly, using the underneath equations (7) and (8), an RNN can be represented,

$$x_i(t) = \sum_j y_j(t) w_{ij}(t) \quad (7)$$

$$y_i(t) = f_i(x_i(t)) \quad (8)$$

Anywhere, y_i and w_{ij} requires the neuron's beginning method i at a time t and optimize weights importance. The beginning function f_i is based on the inputs of the network and context layer inputs.

Step 3: Vector results to determine the hidden node activation function, the function are delivered by sigmoid is given as equation (9),

$$f_i = \frac{1}{1 + e^{(-x_i)}} \quad (9)$$

Here, $i = 1,2$ and the RNN output is $Y^{act} = W_{2i} f_i$ for a single output weight matrix.

Step 4: Ahead of the dissemination of the practice, the release function of each neuron is calculated (10) and (11),

$$y_i(t) = f_i(x_i(t), C_i(t)) \quad (10)$$

$$x_i(t) = \sum_{j \in H} y_j(t) w_{ij} + \sum_{j \in I} x_j(t) w_{ij} + \sum_{j \in C} y_j(t - \tau_{ij}) wij \quad (11)$$

Where, f, H, I and C represents a neuron activation function of hidden layer values, the values of the input neurons, neuron storing data values in the last phase of the network. Then x_j is j^{th} input neuron and τ_{ij} is an integer value referring toin connection with the series of cases of displacement.

Here, the value t represents the back propagation error value for giving the neural calculated

Step 5: The back propagation error is found from the equation (12),

$$E_m = Y^{tar} - Y^{act} \quad (12)$$

To minimize the error value, the weights values of RNN are updated with the help of the CSO algorithm.

Weight Updation using CSO algorithm:

In this section, the weight values of the RNN are enhanced with the help of the CSO algorithm. The CSO is based on cat behavior. Every cat in a CSO has its status and speed. Each cat has two modes, namely the observation and monitoring system and all cats are divided into two groups. One group performs a search, the other groups monitor. Each mode, cats have different updating rules for updating process. The step-by-step process of CSO based weight renewal is explained below;

Step 1: Initialization: Here, the solution is considered cat and weight value prey. Initially, random weights are assigned. Random weight values are considered the initial solution. Weight values vary from 0-1.

Step 2: Fitness calculation: After the solution begins, the fitness of each solution is calculated. Maximum predictive accuracy is considered the best exercise measure. Exercise function 13 is given in equation.

$$Fitness = \max(Accuracy) \quad (13)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

Step 3: Seeking mode: After the exercise calculation, the cats are presented in search mode. Here, cats look around the position, which slightly adjusts their position. In this case, cats seek the optimal opportunity to find prey. The process of search mode is described below;

- Initially, every cat will represent its individual situation five times.

- Then, the copy locations are remembered in a visual memory pool.
- Then, the mutation function is used. For this reason, each value of the SMP is slightly changed.

The aeration mutation-based purification formula is given in Equation 15.

$$Y[i]=Y[i]+\Delta Y[i] \quad (15)$$

Every cat will choose a point in the SMP through the best fit to update its situation.

Step 4: Tracing mode
At the point when felines are in disclosure mode, their conduct is repeated to the particles in the PSO algorithm, and every cat follows the worldwide best situation to refresh the feline to its own speed and position. Be that as it may, felines rapidly access the worldwide best an incentive by having the experience of gathering entire feline populaces. This recharging rule may quicken the merger rate. The refresh equations based on the observation mode are given in Figures 16 and 17.

$$V_j(k+1)=\omega.V_j(k)+C.rand.\left[Y_{best}(k)-Y_j(k)\right] \quad (16)$$

$$Y(k+1)=Y(k)+Y(k+1) \quad (17)$$

Where;

Y_{best} → Great condition of a cat with excellent fitness

Y_j → Position of cat_j

k → Iterations

C → Constant

$rand$ → The random number of the interval varies the ranges from [0,1]

Step 5: Termination criteria
This mechanism runs the maximum number of cycles and stops the operation only when the best exercise value is selected, and this best exercise-based weight value is allocated to the RNN classification.

Testing stage
After the training process, we do the testing process. Here, first, we present the Testing set (TS) to the pre-processing stage. The processed data is presented before the dimensional reduction process. Then, the reduced dataset is churn or non- churn depend on the threshold value to validate the data provided for the RNN classification. Data is a source of data if the score value is above the threshold, otherwise the data is not a source. The prediction condition is given in Equation (16)

$$Decision=\begin{cases} T_h \geq score\,;\, data\ is\ Non-churn \\ T_h < score\,;\, data\ is\ churn \end{cases} \quad (16)$$

RESULT AND DISCUSSION
This subdivisioncovers the conclusion and discussion of the Consumer Sour Prediction Program planned. To implement the proposed method, we utilized MATLAB version (7.12). This proposed mechanism features the Intel Core i5 processor on Windows, through a speed of 1.6 GHz and 4GB of RAM. The proposed system will be publicly available on the Internet have been tested in clinical datasets plant leaf.

Dataset Description
In this investigation, we use the Consumer churn data set developed after the **Tera data Center** at Duke University in the United States [32]. In the United States (US) these customers generated datasets from July to December 2001. These current data are recognized as calibration data. In this essential dataset, 100,000 entries of consumer data are made up of data, and the remaining 50% are uncomplicated. The size of the indicator properties in that data set is 171

Evaluation Matrices
The performance of the proposed methodology is evaluated in terms of different metrics namely, Sensitivity, Specificity, Accuracy, PPV, NPV, FPR, FNR, and FDR.

Sensitivity
The sum of true positives and false negatives sensitive to the ratio of the number of true positives is known as sensitivity.

$$Sensitivity=\frac{No.of\,(T^P)}{No.of\,(T^P)+No.of\,(F^N)}\times 100 \quad (1)$$

Specificity: Specific negativity is the ratio of the sum of one true negative and false positive.

$$Specificity=\frac{No.of\,(T^N)}{No.of\,(T^N)+No.of\,(F^P)}\times 100 \quad (2)$$

Accuracy: Sensitivity and specificity are calculated by accuracy. This is indicated as follows,

$$Accuracy=\frac{T^P+T^N}{T^P+T^N+F^P+F^N}\times 100 \quad (3)$$

Positive Predictive Value (PPV): The positive predictive value of the test is considered positive effects of the fraction:

$$PPV=\frac{T^P}{T^P+F^P} \quad (4)$$

Negative Predictive Value (NPV): The negative predictive value of the test is deemed negative effects of the fraction:

$$NPV=\frac{T^N}{T^N+F^N} \quad (5)$$

False Positive Rate (FPR): The number of false negative predictions FPR is divided by the total number of opposites. This can also be measured as 1 - specification.

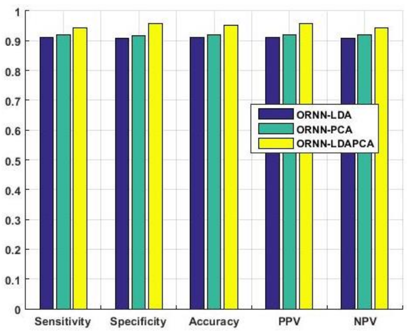
$$FPR=\frac{F^P}{F^P+T^N} \quad (6)$$

False Negative Rate (FNR): FNR is measured as the number of false negative predictions divided by the total number of negatives.

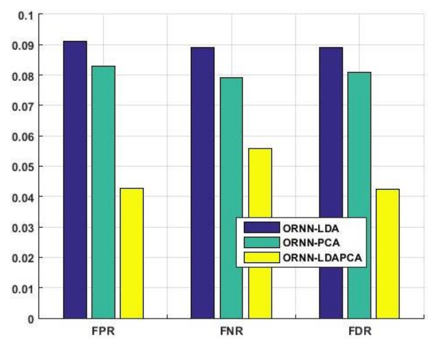
$$FNR = \frac{F^N}{F^N + T^P} \quad (7)$$

Comparative Analysis

The major involvement of our work is to improve an effective churn forecasting model. Since its inception, the input dataset has been improved as a dimension reduction function. For dimensional reduction, the proposed method uses a hybrid technique. Here, PCA and LDA algorithm are hybridized to reduce dimensionality. PCA is one of the standard unsupervised techniques to decrease the dimensionality of the dataset lacking defeated important information in the original dataset. But this may result in lower recognition rates. To overcome these issues in PCA, PCA Strategy is hybridized with LDA. The LDA can perform dimension reduction within the class dispersion phase and dissolve the matrix between classes. LDA shows improved results in a large database[27].



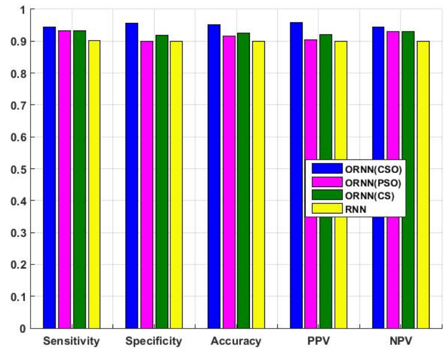
(a)



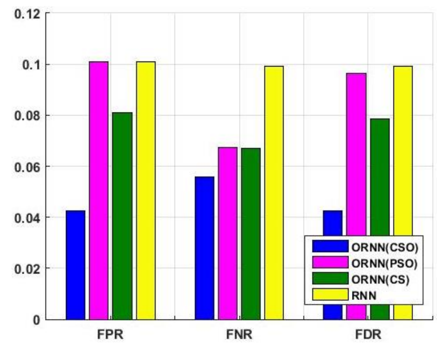
(b)

Figure 3: Comparative analysis for the proposed method based on dimensionality reduction method

Figure 3 (a) above shows the comparative analysis for the proposed ORNN-LDAPCA and the existing ORNN-LDA and ORNN-PCA. To exhibit the prevalence of the proposed half breed PCALDA model over the PCA, LDA and two-arrange model. There is an ideal arrangement among PCA and LDA, which implies that the most straightforward and most prejudicial data can be ensured. Interestingly, the presentation of the 2-level model might be impacted by a lot of visually impaired data inside the initial scarcely any key parts. In this way, a cross breed model joining PCA and LDA is superior to consolidating dimensional decrease by means of a two-arrange framework. The image above shows the examination of affectability, particularity, exactness, PPV, NPV, FPR, FNR and FDR esteems. Likewise, the proposed crossover PCA-LDA model can be considered as an extraordinary instance of grouping creation segregation models. So PCA can be replaced by other production models and LDA with other discrimination protocols.



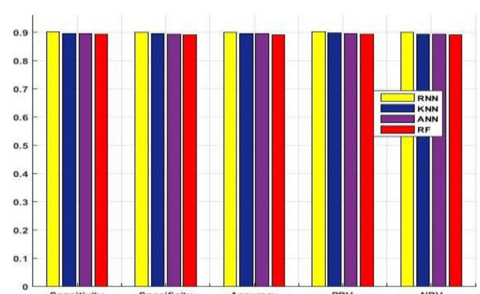
(a)



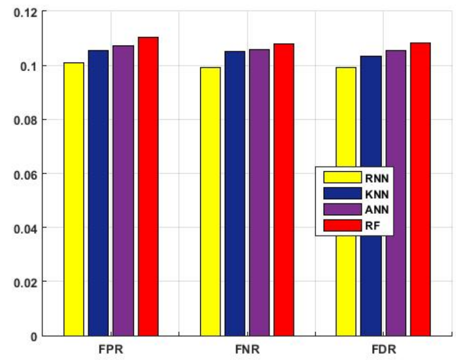
(b)

Figure 4: Comparative analysis for the proposed method based on a different optimization algorithm

To progress the presentation of the RNN classifier, the CSO algorithm is used in this paper. For comparison purposes, in this paper, particle swarm optimization (PSO), crane search optimization (CS), and without optimization are used. The proposed CSO algorithm is similar to the PSO, but the convergence speed of the CSO is better than that of the PSO. CS behavior is based on the behavior of cookies. This algorithm easily falls into the local Optima. Therefore, in this paper, the CSO algorithm is used. When analyzing Figure 4 (a), our proposed prediction method achieves a maximum accuracy of 95.08%, which is 91.59% for using RNN + PSO, 92.61% for using RNN + CS, and 89.99% for using RNN + CS based forecasting . Due to the hybrid dimension reduction method and the optimal classifier, our proposed technique accomplishes improved accuracy compared to other methods. Similarly, our proposed method achieves better outputs for sensitivity, specification, PPV and NPV. When analyzing Figure 4 (b), our proposed method achieves the minimum error value compared to other methods.



(a)



(b)

Figure 5: Comparative analysis for the proposed method based on different classifiers

In Figure 5, the performance of the proposed method is analyzed based on different measurements with different classifiers. When analyzing Figure 5 (a), our proposed method achieves a maximum of 90.5% accuracy, which is 89.4% for using KNN-based prediction, 89.3% for using ANN-based prediction, and 89% for RF-based prediction. Similarly, our proposed method has a maximum of 90.08%, a specificity of 89.89%, a PPV of 0.900693756 and an NPV of 0.899091826. When analyzing Figure 5 (b), our proposed approach has at least 0.101089588, FNR 0.099127676 and FDR 0.099306244 compared to other classifiers. Results from the proposed method achieves better results when compared to other approaches is clearly understood. This is due to hybrid dimension reduction and optimal RNN classification.

CONCLUSION:

This study is based on the CSP hybrid dimension reduction approach and optimal RNN. Here, initially, the dimensionality of the dataset is reduced with the help of the PCLDA algorithm, which is a hybrid of PCA and LDA. Then, the reduced dataset is handed over to the ORNN. RNN classifier performance is improved with the benefit of CSO algorithm. Finally, depend on the classifier, the data is classified as Churn or non-Churn data. The effectiveness of the proposed method has been calculated with the help of different measurements. All the results presented in the Results section are clearly shown, and our proposed method achieved better performance compared to different methods.

REFERENCE

1. Qureshi, S. A., Rehman, A. S., Qamar, A. M., Kamal, A., & Rehman, A. (2013). Telecommunication subscribers' churn prediction model using machine learning. Eighth International Conference on Digital Information Management (ICDIM 2013).
2. Amin, A., Al-Obeidat, F., Shah, B., Tae, M. A., Khan, C., Durrani, H. U. R., & Anwar, S. (2017). Just-in-time customers churn prediction in the telecommunication sector. The Journal of Supercomputing.
3. Shirazi, F., & Mohammadi, M. (2018). A big data analytics model for customer churn prediction in the retiree segment. International Journal of Information Management.
4. P, B., & S, N. G. (2017). A Review on Churn Prediction Modeling in Telecom Environment. 2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS).
5. Alamsyah, A., & Salma, N. (2018). A Comparative Study of Employee Churn Prediction Model. 2018 4th International Conference on Science and Technology (ICST).
6. Ahmed, A., & Linen, D. M. (2017). A review and analysis of churn prediction methods for customer retention in telecom industries. 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS).
7. Mishra, A., & Reddy, U. S. (2017). A comparative study of customer churn prediction in telecom industry using ensemble based classifiers. 2017 International Conference on Inventive Computing and Informatics (ICICI).
8. Dolatabadi, S. H., & Keynia, F. (2017). Designing of customer and employee churn prediction model based on data mining method and neural predictor. 2017 2nd International Conference on Computer and Communication Systems (ICCCS).
9. Chu, C., Xu, G., Brownlow, J., & Fu, B. (2016). Deployment of churn prediction model in financial services industry. 2016 International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC).
10. Spiteri, M., & Azzopardi, G. (2018). Customer Churn Prediction for a Motor Insurance Company. 2018 Thirteenth International Conference on Digital Information Management (ICDIM).
11. Runge, J., Gao, P., Garcin, F., & Faltings, B. (2014). Churn prediction for high-value players in casual social games. 2014 IEEE Conference on Computational Intelligence and Games.
12. Xia, G., Wang, H., & Jiang, Y. (2016). Application of customer churn prediction based on weighted selective ensembles. 2016 3rd International Conference on Systems and Informatics (ICSAI).
13. Shen, Q., Li, H., Liao, Q., Zhang, W., & Kalilou, K. (2014). Improving churn prediction in telecommunications using complementary fusion of multilayer features based on factorization and construction. The 26th Chinese Control and Decision Conference (2014 CCDC).
14. Halim, J., & Vucetic, J. (2015). Increasing Effectiveness of Churn Prediction Software. 2015 Annual Global Online Conference on Information and Computer Technology (GOCICT).
15. Xia, X., Zeng, L., & Yu, R. (2018). HMM of Telecommunication Big Data for Consumer Churn Prediction. 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of

People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IO

16. Hammoudeh, A., Fraihat, M., & Almomani, M. (2019). Selective Ensemble Model for Telecom Churn Prediction. 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT).
17. Do, D., Huynh, P., Vo, P., & Vu, T. (2017). Customer churn prediction in an internet service provider. 2017 IEEE International Conference on Big Data (Big Data).
18. Verbeke, W., Martens, D. and Baesens, B., 2014. Social network analysis for customer churn prediction. Applied Soft Computing, 14, pp.431-446.
19. Stripling, E., vanden Broucke, S., Antonio, K., Baesens, B., & Snoeck, M. (2018). Profit maximizing logistic model for customer churn prediction using genetic algorithms. Swarm and Evolutionary Computation, 40, 116–130.
20. Milošević, M., Živić, N., & Andjelković, I. (2017). Early churn prediction with personalized targeting in mobile social games. Expert Systems with Applications, 83, 326–332.
21. Höppner, S., Stripling, E., Baesens, B., Broucke, S. Vanden, & Verdonck, T. (2018). Profit-driven decision trees for churn prediction. European Journal of Operational Research.
22. Lu, N., Lin, H., Lu, J., & Zhang, G. (2014). A Customer Churn Prediction Model in Telecom Industry Using Boosting. IEEE Transactions on Industrial Informatics, 10(2), 1659–1665.
23. Idris, A., & Khan, A. (2016). Churn Prediction System for Telecom using Filter–Wrapper and Ensemble Classification. The Computer Journal, bxv123.
24. Castro, E. G., & Tsuzuki, M. S. G. (2015). Churn Prediction in Online Games Using Players' Login Records: A Frequency Analysis Approach. IEEE Transactions on Computational Intelligence and AI in Games, 7(3), 255–265.
25. Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2018). Customer churn prediction in the telecommunication industry using data certainty. Journal of Business Research.
26. D. Senthil, G. Suseendran" "Data Mining Techniques Using Time Series Research", International Journal of Recent Technology and Engineering, Vol.8,(2S11), September 2019, pp. 121-129.
27. B.Mahalakshmi, G.Suseendran, "A Hybrid Cryptographic Algorithm for Securing Data in Cloud Storage", Journal of Advanced Research in Dynamical and Control Systems, Vol.11(6), July, 2019 pp.695-704.